Chapter 3: Failure Modes and Mechanisms

B. Stark

A critical part of understanding the reliability of any system comes from understanding the possible ways in which the system may fail. In MEMS, there are several failure mechanisms that have been found to be the primary sources of failure within devices. In comparison to electronic circuits, these failure mechanisms are not well understood nor easy to accelerate for life testing. In any discussion of failures, the definition of failure mechanisms, or causes of failure, often overlaps with the definition of failure modes, or observable failure events. To alleviate this confusion this chapter has been roughly organized by failure modes, with mechanisms being described within the sections on the modes they cause. Failure mechanisms that do not have clearly associated modes are discussed at the end of this chapter.

I. Mechanical Fracture

Mechanical fracture is defined as the breaking of a uniform material into two separate sections. In MEMS it will usually lead to the catastrophic failure of a device, although there are some structures that will have more moderate performance degradations.[5,8] No matter what the actual outcome, any fracturing is a serious reliability concern. There are three types of fractures, ductile, brittle, and intercrystalline fracture. Ductile fracture, as the name implies, occurs in ductile materials. It is characterized by almost uninterrupted plastic deformation of a material. It is usually signified by the necking, or extreme thinning, of a material at one specific point. Brittle fracture occurs along crystal planes and develops rapidly with little deformation. Intercrystalline fracture is a brittle fracture that occurs along grain boundaries in polycrystalline materials, often beginning at a point where impurities or precipitates accumulate. For MEMS the latter two types of fracture are more common. To understand the actual causes of fracture and the methods for predicting it, several terms must be first defined.[27]

A. Definitions

Mechanical failure in a crystal lattice occurs when an applied stress exceeds the failure stress of the structure. Stresses are separated into the two categories of normal and shear stress. Normal stress is defined as stress perpendicular to a plane in a material, while shear stress occurs parallel to a plane, as shown in Figure 3-1. In solid materials, stress is linearly related to a concept called strain, which is the fractional elongation of a material. The proportionality constant between stress and strain is, for small normal loads, called the modulus of elasticity, or Young's modulus. The actual deformation of a cubic volume will depend upon all the stresses applied to it:

$$
\begin{bmatrix} s_{xx} \\ s_{yy} \\ s_{zz} \\ t_{xy} \\ t_{xz} \\ t_{yz} \end{bmatrix} =
\begin{bmatrix}
E_{11} & E_{12} & E_{13} & E_{14} & E_{15} & E_{16} \\
E_{21} & E_{22} & E_{23} & E_{24} & E_{25} & E_{26} \\
E_{31} & E_{32} & E_{33} & E_{34} & E_{35} & E_{36} \\
E_{41} & E_{42} & E_{43} & E_{44} & E_{45} & E_{46} \\
E_{51} & E_{52} & E_{53} & E_{54} & E_{55} & E_{56} \\
E_{61} & E_{62} & E_{63} & E_{64} & E_{65} & E_{66}
\end{bmatrix}
\begin{bmatrix} e_{xx} \\ e_{yy} \\ e_{zz} \\ e_{xy} \\ e_{xz} \\ e_{yz} \end{bmatrix} \tag{3-1a}
$$

where

σ = normal stress

τ = shear stress

ε = strain

E = Elastic Modulus

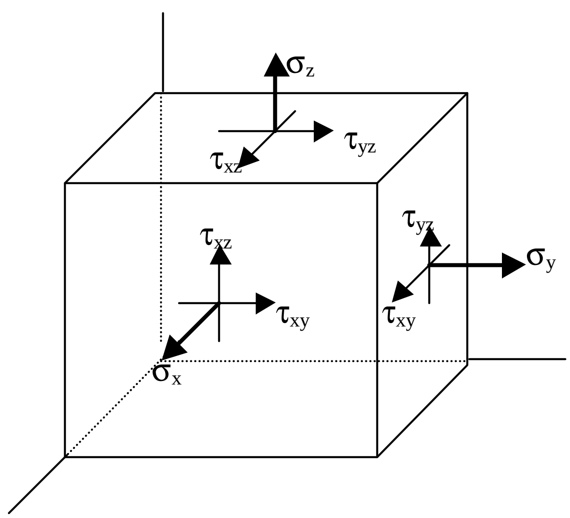

Figure 3-1: Generalized stress states on a 3-D unit cube.

One important aspect of this tensor is that $E_{ij}=E_{ji}$, so that there are actually only 21 independent constants. Further simplifying this effect is the internal symmetry of most crystals. In cubic crystals, such as Si and GaAs, the tensor reduces to:

$$
\begin{bmatrix} s_{xx} \\ s_{yy} \\ s_{zz} \\ t_{xy} \\ t_{xz} \\ t_{yz} \end{bmatrix} = \begin{bmatrix} E_{11} & E_{12} & E_{12} & 0 & 0 & 0 \\ E_{12} & E_{11} & E_{12} & 0 & 0 & 0 \\ E_{12} & E_{12} & E_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & E_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & E_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & E_{44} \end{bmatrix} \begin{bmatrix} e_{xx} \\ e_{yy} \\ e_{zz} \\ e_{xy} \\ e_{xz} \\ e_{yz} \end{bmatrix} \qquad (3\text{-}1b)
$$

While not all materials have just three independent elastic constants, it is unlikely to find even highly anisotropic crystals with more than nine elastic constants.

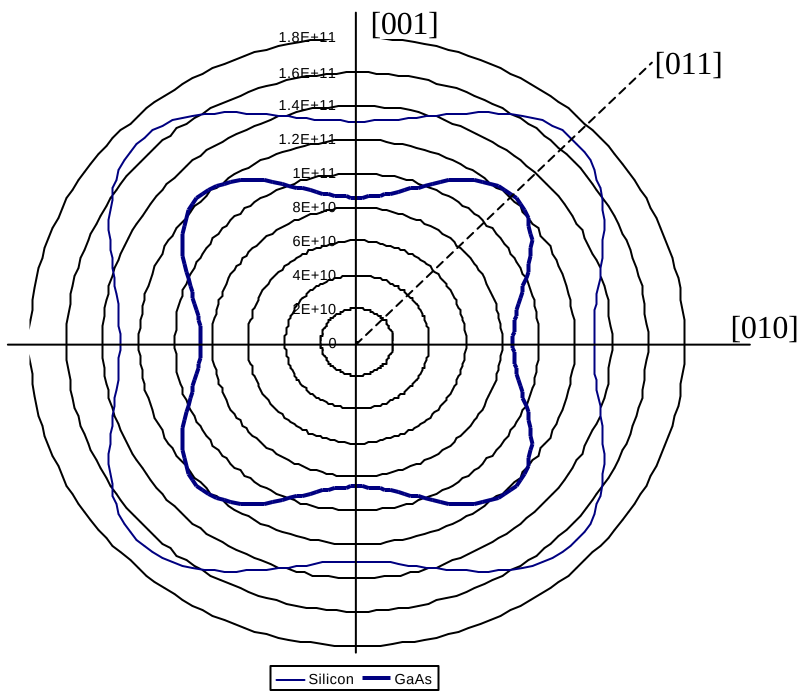

Figure 3-2: Young's modulus as a function of crystalline orientation for Si and GaAs along the <100> axis.

One of the difficulties in using Equation 3-1b is that, in an anisotropic crystal, the elastic modulus will vary with crystalline orientation. To account for this variation a plane modulus is defined with the crystal orientation by

$$
E_{[hkl]} = \frac{s_{[hkl]}}{e_{[hkl]}} \qquad (3\text{-}2)
$$

In order to relate the modulus of a crystalline plane to the elastic constants in Equations 3-1, the following equation is used:

$$E^{-1}(q, f, y) = s_{11} - 2(s_{11} - s_{12} - 0.5s_{44})(l_1^2 l_2^2 + l_2^2 l_3^2 + l_1^2 l_3^2)) \qquad (3\text{-}3)$$

where

$$s_{11} = \frac{E_{11} + E_{12}}{(E_{11} + 2E_{12})(E_{11} - E_{12})}$$

$$s_{12} = \frac{-E_{12}}{(E_{11} + 2E_{12})(E_{11} - E_{12})}$$

$$s_{44} = \frac{1}{E_{44}}$$

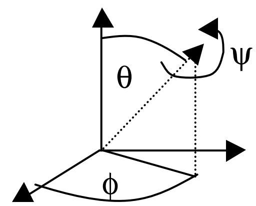

Figure 3-3: Euler's angles.[91] These are the angles formed between the <100> axis and an arbitrary <hkl> axis.

θ, φ, ψ = Euler's angles, defined in Figure 3-2

l_1, l_2, l_3 = the direction cosines defined by the following matrix

$$\begin{bmatrix} l_1 & m_1 & n_1 \\ l_2 & m_2 & n_2 \\ l_3 & m_3 & n_3 \end{bmatrix} = \begin{bmatrix} \cos(q)\cos(f)\cos(y) - \sin(f)\sin(y) & \cos(q)\sin(f)\cos(f) - \sin(f)\sin(y) & -\sin(q)\cos(y) \\ -\cos(q)\cos(f)\sin(y) - \sin(f)\sin(y) & -\cos(q)\sin(f)\cos(f) - \sin(f)\sin(y) & \sin(q)\sin(y) \\ \cos(q)\cos(f) & \cos(q)\cos(f) & \cos(q) \end{bmatrix}$$

The modulus of Si and GaAs as a function of crystalline orientation is shown in Figure 3-2.[32]

In common nomenclature the constants, s_{11}, s_{12}, and s_{44} are called compliance coefficients. This equation reveals that the {100} planes of Si have an elastic modulus of 130 GPa, while the {110} planes have a modulus of 165 GPa. For the {111} planes, with a θ and φ angle of 45° and a ψ angle of 0°, the value of $E_{\{111\}}$ is 187 GPa, which is the stiffest plane in silicon. As a result, wear effects will be most severe in the[100] direction because it has the lowest stiffness of any crystal planes in silicon. It must also be noted that E_{11}, E_{12}, and E_{44} are usually defined relative to the <110> planes, while s_{11}, s_{12}, and s_{44} are generally defined relative to the <100> planes.

Poisson's ratio is also orientation-dependent, with the basis vector given along with the value of the number. Poisson's ratio is normally defined in terms of the elastic compliance coefficients as $\nu = -s_{12}/s_{11}$. If a longitudinal stress is considered in a direction that is displaced from the {100} planes by angles θ, φ, and ψ, it has been proven [48,49] that

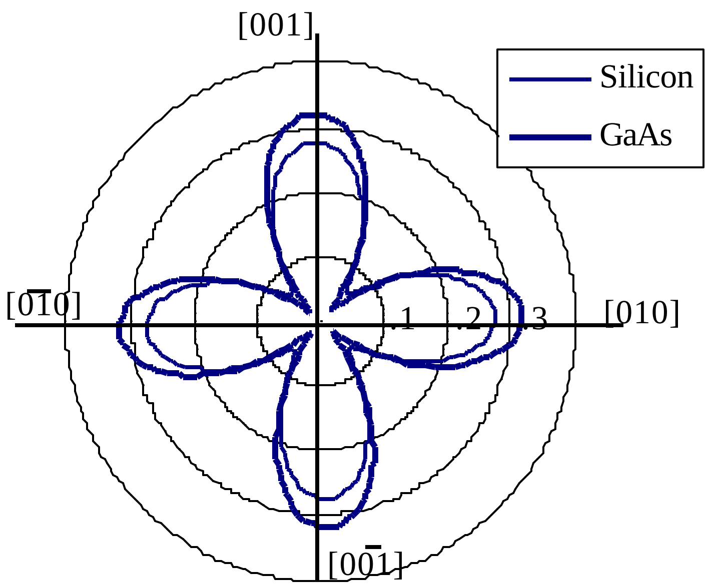

Figure 3-4: Poisson's ratio as a function of angle in the (100) plane with l and m varying in the (100) plane.

$$n = -\frac{s_{12} + (s_{11} - s_{12} - .5s_{44})(l_1^2 m_1^2 + l_2^2 m_2^2 + l_3^2 m_3^2)}{s_{11} - 2(s_{11} - s_{12} - .5s_{44})(l_1^2 l_1^2 + l_2^2 l_3^2 + l_1^2 l_3^2)} \tag{3-4}$$

While these considerations are important in the study of MEMS, they are difficult to resolve analytically. For simplicity's sake, researchers design structures that will only be forced in orthogonal directions, so that Young's modulus and Poisson's ratio can be treated as uniform values. For this reason, the remainder of this guideline will treat Young's Modulus and Poisson's ratio as a single value, with the implicit understanding that these quantities are actually dependent upon crystal structure.

Once the definitions of stress and strain are understood, it is possible to understand how stress leads to failure.

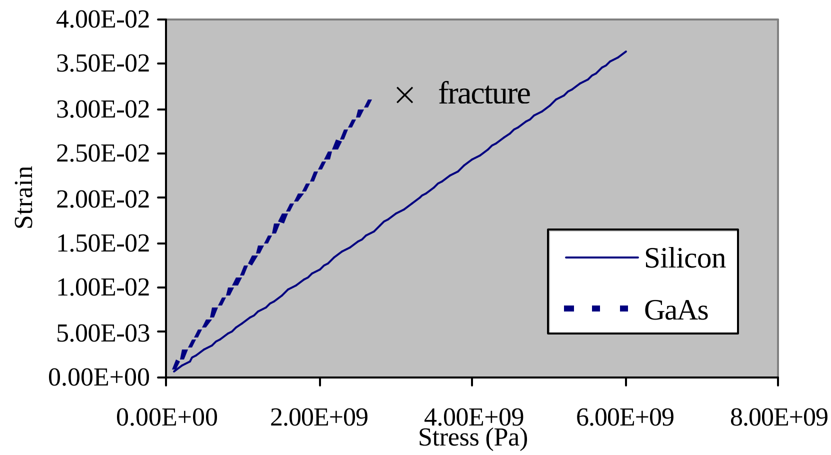

Figure 3-5: Stress versus strain relationships for bulk Si and GaAs[1].

B. Stress-Induced Failure

In Figure 3-2, the process of crystal lattice failure is illustrated through a diagram of stress versus strain. As can be seen, the application of stress causes a linear increase in strain until fracture. This is a function of the brittle properties of these materials; brittle materials deform elastically until fracture occurs. To understand the fracture tolerances of a MEMS device, as in any mechanical structure, one needs to determine the maximum stresses.

The maximum stress in a device usually occurs near stress risers, or concentrators. Stress concentration occurs when there is a sudden change in the cross section of a material. At these points, stress is usually non-uniformly distributed and somewhat difficult to analytically resolve. Since most engineers are more concerned with maximum stress rather than average stress, this value can be calculated by defining the stress concentration factor K as:

$$K = \frac{\sigma_{\max}}{\sigma_{ave}} \qquad (3\text{-}5)$$

---

[1] This chart is idealized. In actuality there is a small curvature to the stress strain curve of any material.

where

σ_{max} = maximum stress at a stress concentration point

σ_{ave} = average stress at a stress concentration point

K is a function of the geometry of the stress riser and is typically graphically represented. σ_{ave} can be calculated from basic structural analysis, which allows σ_{max} to be determined.

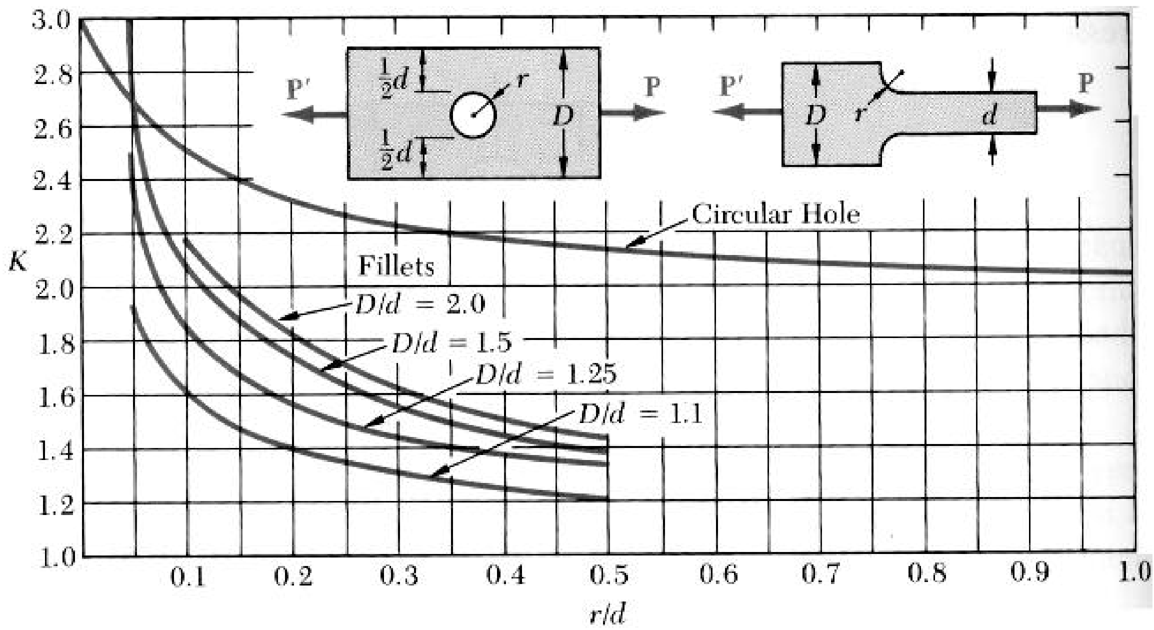

Figure 3-6: Stress concentration factor as a function of geometry for thin beams. (from [11])

While analysis of stress will determine if a material has exceeded its fracture strength, some new concepts will have to be introduced to understand what factors limit the strength of materials. Many of the mechanical failures in crystalline solids occur as the result of defects in crystal structures. These defects are the result of imperfect techniques in crystal growth and are critically important to the study of the properties of crystals. While a modern silicon wafer will have relatively few defects,[46] other materials used in MEMS have significant defect densities. There are several kinds of defects that need to be examined.

C. Point Defects

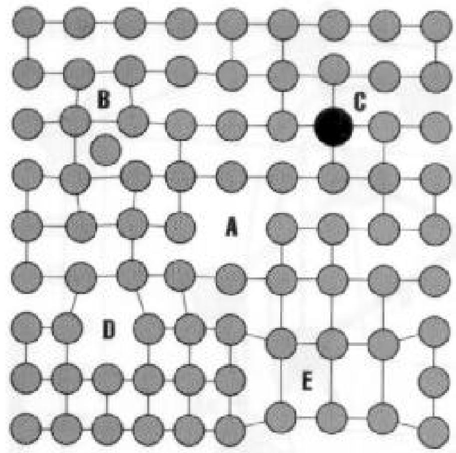

Figure 3-7: Different point defects. A is a vacancy, B is an interstitial. C is a Point replacement while D and E are 2-D mappings of an edge and screw dislocation, respectively.

Point defects are faults at a single point in a crystal. They tend to create very localized internal strains and do not usually have the magnitude of an impact upon crystal lattice integrity that is common to dislocations, which are discussed in the following section. Point defects are categorized into the following groups:

i) Vacancies

A vacancy is the lack of an atom at a specific point in a lattice where one would otherwise be expected. This has the result, as most defects do, of limiting electron mobility. More important for MEMS is the fact that a vacancy will lower the yield strength of the material, as it weakens the lattice strength. The missing atom causes the lattice to compress around the vacancy, which creates an internal stress field that is described by Equation 3-5.

$$s_{rr} = -2s_{jj} = -2s_{qq} = \frac{g_0}{p}\frac{1-2n}{1-n}\frac{1}{r^3} \qquad (3\text{-}6)$$

where

$$g_0 = 2G\Delta V\frac{1-n}{1-2n}$$

G = the shear modulus of a material (E_{44} for cubic crystals)

ΔV = the change in the volume of the solid due to the vacancy.

ii) Interstitial

An interstitial is an additional atom which has become wedged between the atoms of a lattice. An interstitial will have, to a first order approximation, the same stress field as a vacancy, except that the strength factor, g_0, becomes

$$g_0 = \frac{8pG}{3}\frac{1+n}{1-2n}hr_0^3 \qquad (3\text{-}7)$$

where

$$h \approx \frac{r_0}{A} - 1$$

r_0 = the radius of the foreign atom

A = the lattice parameter

iii) Point Replacement

In this case a single atom has been replaced by an atom of a different element. Often this is done intentionally for doping purposes, but sometimes it occurs accidentally as the result of disorder in the lattice or impurities in the melt. Theses defects will have differing effects on the mechanical properties of solids, but will usually not be as large in magnitude as vacancies or interstitials.[53]

D. Dislocations

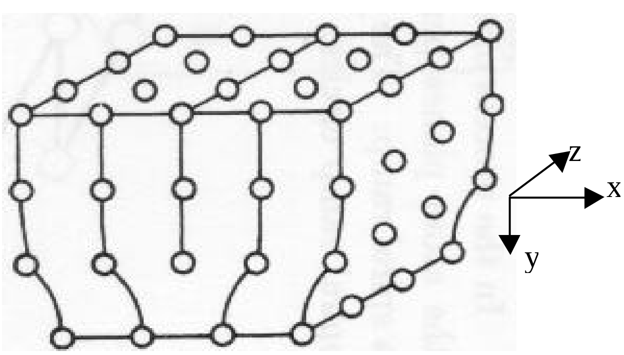

Figure 3-8a: Edge dislocation. (from [37])

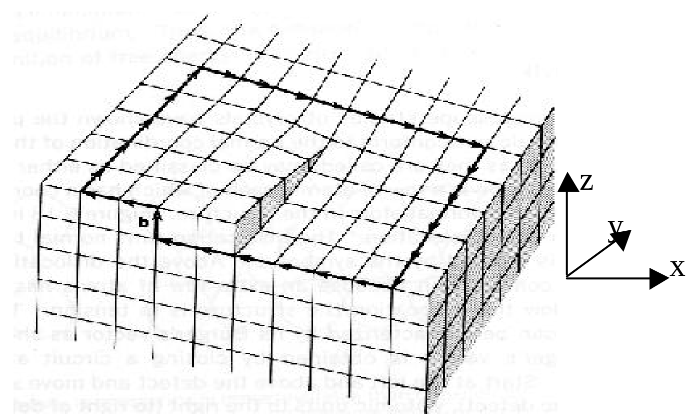

Figure 3-8b: Screw dislocation. (from [37])

A dislocation is a one-dimensional array of point defects in an otherwise perfect crystal. They occur when a crystal is subjected to stresses that exceed the elastic limits of materials. Dislocations are usually introduced into a crystal through the presence of a temperature gradient during crystal growth. Modern wafer processing techniques produce extremely low dislocation densities on wafers. Dislocations can be separated into two types:

i) Edge Dislocation

In this case a whole row of atoms is out of phase with respect to the rest of the lattice, as shown in Figure 3-8a. The result of this phenomenon is a physical barrier in the crystal that scatters electrons and weakens the crystal. An edge dislocation creates a stress field that is defined by Equations 3-8a-e:

$$s_x = \frac{Gb}{2p(1-n)} \frac{y(3x^2+y^2)}{(x^2+y^2)^2} \quad (3\text{-}8a)$$

$$s_y = \frac{Gb}{2p(1-n)} \frac{y(x^2-y^2)}{(x^2+y^2)^2} \quad (3\text{-}8b)$$

$$s_z = n(s_x + s_y) \quad (3\text{-}8c)$$

$$t_{xy} = -\frac{Gb}{2p(1-n)} \frac{x(x^2-y^2)}{(x^2+y^2)^2} \quad (3\text{-}8d)$$

$$t_{xz} = t_{yz} = 0 \quad (3\text{-}8e)$$

where

x, y, z = distance from edge dislocation, with the dislocation in the plane of x and the z axis is tangent to the dislocation.

b = The magnitude of the Burgers vector of the dislocation ~ a few lattice spacings.

ii) Screw Dislocation

This fault is much the same as an edge dislocation except that it is shaped like a spiral staircase, as shown in Figure 3-8b. Screw dislocations also create stress fields in solids, as defined below:

$$t_{xz} = \frac{Gb}{2p} \frac{y}{x^2+y^2} \quad (3\text{-}9a)$$

$$t_{yz} = -\frac{Gb}{2p}\frac{x}{x^2 + y^2} \tag{3-9b}$$

$$s_x = s_y = s_z = t_{xy} = 0 \tag{3-9c}$$

The essential feature to recognize from these equations is that dislocations introduce stresses internal to materials that will significantly weaken crystal lattices. While these stresses decrease quadratically with distance from the dislocation, there clearly will be a strong local internal stress created by these features. Another factor to consider is that the stress fields from different dislocations will interact, creating internal forces. From a MEMS reliability standpoint, this means that using high quality wafers with smaller numbers of dislocations will ultimately increase device reliability and lifetime. For more information on this subject, Reference [37], “Elementary Dislocation Theory” by Weertman and Weertman offers a good, in-depth discussion of dislocation theory.

E. Precipitates

In metals containing another element in a supersaturated solid solution, this solution tends to precipitate in the form of a compound with the solvent metal.[53] In the presence of a dislocation, atoms will precipitate into the dislocation, which will occur at the rate:

$$n(t) \propto \left(\frac{t}{T}\right)^{2/3} \tag{3-10}$$

where

n(t) = the number of atoms precipitating in a time, t.

T = temperature

This phenomenon is used to harden materials by allowing precipitates to lock dislocation and prevent them from moving through the lattice. While this is found to be useful in construction materials, it creates problems for MEMS devices using metallic compounds, such as GaAs. The formation of precipitates creates an internal stress, which can significantly weaken crystal lattices and is discussed in great detail in Reference [53].

F. Fracture Strength

The impetus for studying defects is that, for brittle materials, the fracture strength is a function of the largest crystal defect. For a defect of length c, the fracture strength can be determined by:[114]

$$s_F = \frac{K_{Ic}}{Y\sqrt{c}} \tag{3-11}$$

where

K_{Ic} = Fracture toughness

σ_f = fracture strength

Y = dimensionless parameter that depends on the geometry of the flaw

It is sometimes useful to approximate the defect as being penny shaped with a radius, c, in which case the fracture strength is:[17]

$$s_F = \frac{1.6K_{Ic}}{\sqrt{pc}} \tag{3-12}$$

Normally, a Gaussian distribution of crystal defects is assumed for a given volume. This would imply that there is also a normal distribution of fracture strengths, as described below:

$$f(s) = (2pd^2)^{-\frac{1}{2}} e^{-\frac{(s-\bar{s})^2}{2d^2}} \tag{3-13}$$

where

d^2 = standard deviation in fracture strength

$\bar{s}$ = mean fracture strength

For this kind of fracture stress distribution, the probability of failure of a body exposed to a stress field is modeled by the Weibull probability distribution function of:

$$P_f = 1 - e^{-\int_v \left(\frac{s - s_u}{s_0}\right)^m dv} \text{ [40]} \tag{3-14}$$

where

V = the volume of the stressed body

σ_u = a lower stress limit which is usually equal to zero in brittle materials (Pa)

σ_0 = a parameter related to the average fracture stress (Pa)

m = the Weibull modulus, a measure of the statistical scatter displayed by fracture events

For simple geometries, Equation 3-14 can be simplified to:

$$\ln\left[\ln\left(\frac{1}{1-P_f}\right)\right] = m \ln s + const \qquad (3\text{-}15)$$

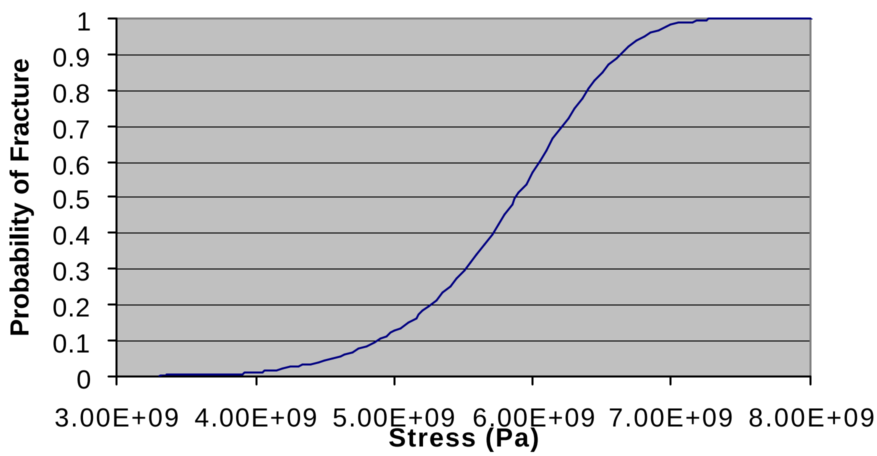

Figure 3-9: Representative plot of the probability of the fracture of silicon under applied stress.

The Weibull modulus, which can be used to measure the reliability of an engineering material, can be determined by a linear curve fit on a simple ln/ln-ln plot. Thus, the probability of fracture will be randomly distributed by a Weibull model, as shown in Figure 3-9. As such, the fracture strength of a material can only be expressed as a probability distribution and not as a specific value, which presents a challenge for reliability engineers. Since it is not known what the ultimate failure strength of a device will be, testing must be done on all devices prior to insertion into the market to eliminate devices with unacceptably low fracture strengths.

For common MEMS materials, several fracture studies have been conducted. The median fracture strength of silicon beams has been reported to be on the order of 6 GPa with a Weibull modulus of 10.1, although these values are strongly dependent upon the finished composition of the beam.[40] GaAs had a fracture strength of less than half that of silicon, at

2.7 GPa. These values compare favorably to construction steel, which fractures under 1 GPa. While this study provides a baseline for the strength of Si and GaAs, they are not universally applicable. Fracture studies need to be conducted on every process, as mechanical properties of materials are highly dependent on processing conditions.

G. Fatigue

Fatigue is a failure mechanism caused by the cyclic loading of a structure below the yield or fracture stress of a material. This loading leads to the formation of surface microcracks that cause the slow weakening of the material over time and create localized plastic deformations. While brittle materials do not experience macroscopic plastic deformation, they will still experience fatigue, albeit in a much longer time frame than in ductile materials.

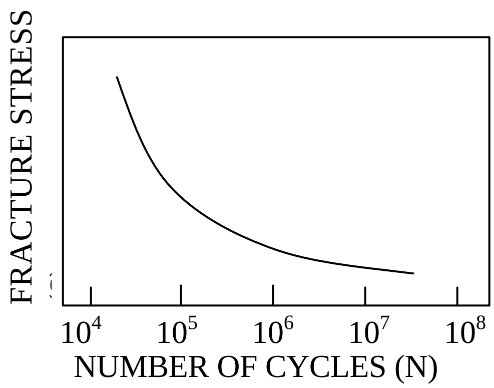

Figure 3-10: Typical SN curve for a ductile material.

Fatigue is typically modeled with a plot known as the SN curve. The plot, shown in Figure 3-11, relates the fracture stress, S, to the number of cycles of loading and unloading a material. As shown in the figure, the fracture stress decreases with time and can eventually fail.

Fatigue also causes a gradual change in the properties of a material. After repeated cycling, which is often on the order of billions of cycles, Young's modulus will gradually shift. This shift will change the resonant frequency of many devices and degrade sensor outputs. Also affected is the dampening coefficient, which will increase over time and change the resonant frequency and Q factor of a structure. Electrical resistance of many structures will also increase over time. The combined effect of these changes can lead to degradation failure. Table 3-1 provides order-of-magnitude estimates in the changes that might be expected in these values over the lifetime of a device.[95]

Property	P+ Silicon		P+ Silicon coated with aluminum		P+ Silicon coated with nitride		P+ Silicon coated with oxide	
	original	aged	original	aged	original	aged	original	aged
Young's Modulus (GPa)	129	136	117	113	134	125	127	123
Dampening Coefficient (Hz)	600	770	741	779	436	588	631	685
Resistance (Ω)	9.9	15.3	19.9	22.7	4.93	9.55	10.3	12.5

Table 3-1: Fatigue induced materials degradations.[95]

While this table is only indicative of one study, it does show that fatigue is a real concern for MEMS.

II. Stiction

One of the biggest problems in MEMS has been designing structures that can withstand surface interactions. This is due to the fact that, when two polished surfaces come into contact, they tend to adhere to one another. While this fact is often unimportant in macroscopic devices, due to their rough surface features and the common use of lubricants, MEMS surfaces are smooth and lubricants create, rather than mitigate, friction.[19] As a result, when two metallic surfaces come into contact, they form strong primary bonds, which joins the surfaces together. This is analogous to grain boundaries within polycrystalline materials, which have been found to be often stronger than the crystal material itself. However, adhesive boundaries are usually not as strong as grain boundaries, due to the fact that the actual area of contact is limited by localized surface roughness and the presence of contaminants, such as gas molecules.

Adhesion is caused by van der Waals forces bonding two clean surfaces together. The van der Waals force is a result of the interaction of instantaneous dipole moments of atoms. If two flat parallel surfaces become separated by less than a characteristic distance of z_0, which is approximately 20 nm, the attractive pressure will be:

$$P_{vdw} = \frac{A}{6pd^3} \text{ [110]} \qquad (3\text{-}16)$$

where

A = Hamaker constant (1.6 eV for Si)

d = the separation between the surfaces

While this equation, since it ignores the repulsive part of surface forces, overestimates the force of adhesion by at least a factor of 2, it is a good order of magnitude approximation for adhesive forces. Typical values of d are on the order of several Ångstroms.

In most MEMS devices, surface contact causes failure. With the noted exception of the small contact areas in microbearings, when surfaces come into contact in MEMS, the van der Waals force is strong enough to irrevocably bond them. Although some devices, such as microswitches, are designed to combat this problem through strong actuator networks, most devices must be designed to eliminate any surface interactions, in order to avoid the effects shown in Figure 3-11.

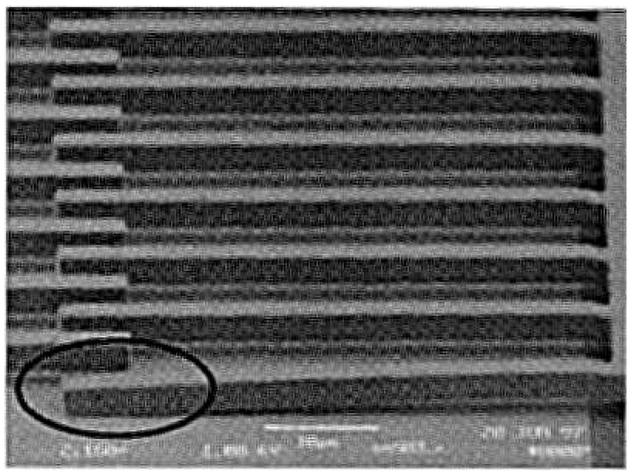

Figure 3-11: Polysilicon cantilever adhering to substrate.

III. Wear

Wear is an event caused by the motion of one surface over another. It is defined as the removal of material from a solid surface as the result of mechanical action.[27] While there are some mechanical operations, such as polishing and sharpening, that utilize wear in a constructive manner, wear is generally considered an undesirable effect in MEMS. There are four main processes that cause wear. They are called adhesion, abrasion, corrosion, and surface fatigue.

Adhesive wear is caused by one surface pulling fragments off of another surface while they are sliding. This is caused by surface forces bonding two materials together. When the bonds break, they are unlikely to separate at the original interface, which fractures one of the materials. The volume of a material fractured by adhesive wear is determined by the relationship:[27]

$$V_{AW} = \frac{k_{AW} F x}{3\sigma_y} \quad (3\text{-}17)$$

where

σ_y = yield strength of the material

k_{AW} = material dependent wear constant

x = sliding distance

F = load on the material

k_{AW} is material dependent. Several useful approximations have been developed for non-metallic wear between different types of materials:

	Nonmetal on identical nonmetal	*Nonmetal on like nonmetal*	*Nonmetal on unlike nonmetal*	*Layer-lattice nonmetal on unlike material*
Unlubricated	1.8×10^{-5}	9×10^{-6}	4.5×10^{-6}	1.5×10^{-6}
Poor lubricant	6×10^{-6}	3×10^{-6}	1.5×10^{-6}	6×10^{-7}
Good lubricant	3×10^{-6}	1.2×10^{-6}	6×10^{-7}	3×10^{-7}
Excellent lubricant	1.5×10^{-6}	6×10^{-7}	3×10^{-7}	9×10^{-8}

Table 3-2: Wear coefficients for nonmetals[1].[120]

As a general rule, adhesive wear will be minimized with dissimilar hard materials.

Initial studies on the long-term effects of adhesive wear have been completed, with some interesting results being discovered. A study at Sandia National Laboratories found that wear related failures had a high initial infant mortality rate, followed by a decreasing failure rate over time. An interesting part of their discovery was that both the lognormal and the Weibull model of failure rates described wear equally well. While this seems a bit counterintuitive, upon close inspection both of these models have fairly similar shapes of probability density and cumulative distribution functions over the ranges in question.

Abrasive wear occurs when a hard, rough surface slides on top of a softer surface and strips away underlying material. While less prevalent in MEMS than adhesive wear, it can occur

[1] Traditionally the factor of 3 is dropped from Equation 3-15 and these values are expressed as 1/3 of the values in this chart.

if particulates get caught in microgears and can tear apart a surface. Also defined by Equation 3-15, the constant k_{AW} for abrasive wear is usually on the order of 10^{-3} to 10^{-6}.

Corrosive wear occurs when two surfaces chemically interact with one another and the sliding process strips away one of the reaction products. This type of wear could cause failure in chemically active MEMS. Certain types of microfluidic systems and biological MEMS are susceptible to corrosive wear. Corrosive wear is dependent upon the chemical reactions involved, but it can be modeled as:

$$h_{CW} = \frac{k_{CW}\, x}{3} \tag{3-18}$$

where

h_{CW}= depth of wear

k_{CW} = corrosive wear constant, on the order of 10^{-4} to 10^{-5}

Surface fatigue wear occurs mostly in rolling applications, such as bearings and gears. It affects highly polished surfaces that roll instead of sliding. Over time, the continued stressing and unstressing of the material under the roller will cause the appearance of fatigue cracks. These cracks then propagate parallel to the surface of a structure, causing material to flake off the surface. Surface fatigue wear tends to generate much larger particles than other wear mechanisms, with flakes as large as 100 nm being common in macroscopic applications.[27]

In many actuator technologies, wear will increase the voltage required to drive a device. Due to the polishing of the contact surfaces caused by wear, the adhesive forces will increase. The increase in adhesion will require larger input signals to drive a device. The increase in drive signal will, in-turn, increase to force, and thus wear, on a structure. As a result, many structures that have contact surfaces prone to wear, will experience a positive feedback loop between wear and driving voltage that will eventually lead to the catastrophic failure of a device. Either the increase in voltage will create a power drain that exceeds the available power to the system or the increase in voltage will decrease the stability of the actuator until stiction occurs.

IV. Delamination

Delamination is a condition that occurs when a materials interface loses its adhesive bond. It can be induced by a number of means, from mask misalignments to particulates on the wafer during processing. It can also arise as the result of fatigue induced by the long term cycling of structures with mismatched coefficients of thermal expansion.

No matter what the actual cause, the effects of delamination can be catastrophic. If the material is still present on the device, it can cause shorting or mechanical impedance.

Furthermore, the loss of mass will alter the mechanical characteristics of a structure. While the exact changes will depend upon the amount of material that has fallen off, recent work[87] has reported shifts of up to twenty five percent in resonant frequency in some devices.

V. Environmentally Induced Failure Mechanisms

In addition to device operation, there are external effects that can also cause failure in MEMS. Many environmental factors can lead to the development of failure modes. As discussed in the next section, environmental failure mechanisms are one of the biggest challenges facing the insertion of MEMS into space.

A. Vibration

Vibration is a large reliability concern in MEMS. Due to the sensitivity and fragile nature of many MEMS, external vibrations can have disastrous implications. Either through inducing surface adhesion or through fracturing device support structures, external vibration can cause failure. Long-term vibration will also contribute to fatigue. For space applications, vibration considerations are important, as devices are subjected to large vibrations in the launch process.

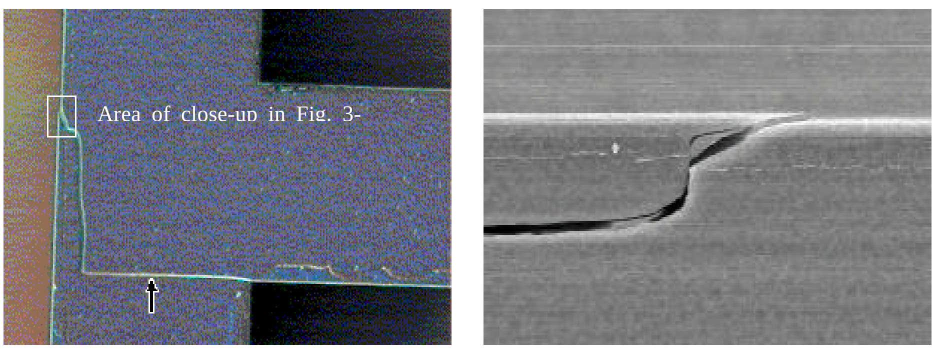

Figure 3-12 (a, b): Cracks in single crystal silicon support beams caused by vibrations from a launch test.[5]

B. Shock

Shock differs from vibration in that shock is a single mechanical impact instead of a rhythmic event. Shock creates a direct transfer of mechanical energy across the device. Shock can lead to both adhesion and fracture. Shock can also cause wire bond shearing, a failure mode common to all semiconductor devices.

C. Humidity Effects

Humidity can be another serious concern for MEMS. Surface micromachined devices, for reasons related to processing, are extremely hydrophilic. In the presence of humidity, water will condense into small cracks and pores on the surface of these structures. When the surface equalizes with the atmosphere there will be a curved meniscus of liquid on the surface, with the two radii of curvature of the meniscus, r_1 and r_2, determined by the expression:

$$\left(\frac{1}{r_1}+\frac{1}{r_2}\right)^{-1} = \frac{gv}{RT\ \log\left(P/P^{sat}\right)} \qquad (3\text{-}19)$$

where

γ = surface tension

v = molar volume

P/P^{sat} = relative vapor pressure of water in the atmosphere

R = gas constant (1.98719 cal/mol-K)

T = temperature

Recent work has shown that condensation on surface micromachined surfaces will lead to an increase in residual stress in the structures.[28] For two surfaces that come into close proximity with each other, the condensation will also create a capillary pressure between the surfaces equal to:

$$P_{cap} = \frac{4gr_1 \cos^2(q)}{d^2} \qquad (3\text{-}20)$$

where

d = the separation between the two plates

θ = contact angle between the surfaces and the liquid

Thus, a hydrophilic surface in a humid atmosphere will experience both condensation, which will create bending moment in structures, and capillary forces, which will create stronger adhesive bonds than Van der Waals forces alone.[110,114]

D. Radiation Effects

While still in its infancy, the field of radiation effects on MEMS is becoming increasingly important. It has long been known that electrical systems are susceptible to radiation and recent research has raised the possibility that mechanical devices may also be prone to radiation-induced damage. Especially sensitive to radiation will be devices that have mechanical motion governed by electric fields across insulators, such as electrostatically positioned cantilever beams. Since insulators can fail under single event dielectric rupture, there is a distinct possibility that these devices will have decreased performance in the space environment. A further complication is the fact that radiation can cause bulk lattice damage and make materials more susceptible to fracture.

Recent work indicates that dielectric layers will trap charged particles, creating a permanent electric field. This permanent field will change resonant characteristics and alter the output of many sensors.[122,123] This may indicate that radiation-tolerant designs will have to limit the use of dielectrics, which could be a challenging design problem. One radiation issue that has received some notice without generating a lot of research is the impact of large atomic mass particles on MEMS. It is known that high Z radiation can lead to fracturing by creating massive disorder within the crystal lattice. Since this radiation source is common in the space environment, it needs to be investigated before MEMS to launch into space.

On one of the few radiation tests to date, one group of surface micromachined devices exposed to gamma ray doses in excess of 25 krad had severe performance degradation. While certainly this reveals no substantive information about the overall radiation tolerance of MEMS as a technology, it raises the possibility that radiation can cause failure within these devices. While more studies need to be conducted before any conclusions can be made, it is important to understand that radiation effects in MEMS is a non-trivial issue that has yet to be fully addressed.[13]

E. Particulates

Particulates are fine particles, that are prevalent in the atmosphere. These particles have been known to electrically short out MEMS and can also induce stiction. While these particles are normally filtered out of the clean room environment, many MEMS are designed to operate outside the confines of the clean room and without the safety of a hermetically sealed package. As a result, devices must be analyzed to ensure that they are particle-tolerant before they can be used as high-rel devices in environments with high particulate densities.

Another area in which contaminants cause problems is in adhesion. Proper device processing requires most materials interfaces to be clean in order to have good adhesion. If dust particles are present, then the two materials will be weakly bonded and are more likely to have delamination problems.

F. Temperature Changes

Temperature changes are a serious concern for MEMS. Internal stresses in devices are extremely temperature dependent. The temperature range in which a device will operate within acceptable parameters is determined by the coefficient of linear expansion. In devices where the coefficients are poorly matched, there will be a low tolerance for thermal variations. Since future space missions anticipate temperatures in the range of –100 to 150^{o}C, thermal changes are a growing concern in MEMS qualification efforts.

Beyond these issues, there are other difficulties caused by temperature fluctuations. Thermal effects cause problems in metal packaging, as the thermal coefficient of expansion of metals can be greater than ten times that of silicon. For these packages, special isolation techniques have to be developed to prevent the package expansion from fracturing the substrate of the device.

Another area that has yet to be fully examined is the effect of thermal changes upon the mechanical properties of semiconductors. It has long been known that Young's modulus is a temperature-dependent value. While it is more or less locally constant for a terrestrial operating range, it may vary significantly for the temperature ranges seen in the aerospace environment.

G. Electrostatic Discharge

Electrostatic discharge, or ESD, occurs when a device is improperly handled. A human body routinely develops an electric potential in excess of 1000V. Upon contacting an electronic device, this build-up will discharge, which will create a large potential difference across the device. The effect is known to have catastrophic effects in circuits and could have similar effects in MEMS. While the effects of ESD on MEMS structures have not been published to date, it can be assumed that certain electrostatically actuated devices will be susceptible to ESD damage.

VI. Stray Stresses

Stray stresses are a failure mechanism that are endemic to thin film structures. Stray stresses are defined as stresses in films that are present in the absence of external forces. In MEMS small stresses will cause noise in sensor outputs and large stresses will lead to mechanical deformation.

Thermal and residual stresses are the two sources of stray stress in MEMS. Thermal stress is a process-induced factor caused by bimetallic warping. Thin films are grown at high temperature and, in the process of cooling to ambient temperature, they contract. While these effects are desirable for the thermocouples, they can cause problem in common MEMS devices. Thermal strains on the order of 5×10^{-4} are commonly observable in MEMS. Residual stresses are a result of the energy configuration of thin films. Caused by the fact that these films are not in their lowest energy state, residual stress can either shrink or expand the substrate. While there are high-temperature techniques for annealing out residual stress, these methods are not always compatible with MEMS processing.

In bulk micromachined devices, stray stresses can cause device warping, while the effects can be much more serious in surface micromachined devices. Since they are made entirely of thin films, surface micromachined devices are exceptionally sensitive to stray stresses. Large residual stresses can cause warping and even the fracturing of the structural material. For this reason, the materials used in surface micromachining are often selected for their ability to be grown with little stray stress.

VII. Parasitic Capacitance

Parasitic capacitance is another failure mechanism in MEMS. Parasitic capacitance does not cause failure in and of itself, but it can be a contributing factor to failure. Parasitic capacitance is defined as an unwanted capacitive effect in a device. While parasitics are unavoidable in devices, they must be minimized for devices to work properly. Parasitic capacitance can cause unwanted electrical and mechanical behavior in devices.

The most common source of parasitics in MEMS is between a device and the substrate. Most MEMS devices consist of a conducting device suspended over a conducting substrate. These two devices have a capacitance between them that is inversely proportional to the distance separating them. This capacitance will exert a force upon the device, creating non-planar displacement and a current flow through the substrate. While some devices use this effect for non-planar actuation, often parasitic capacitance will impinge device performance by causing unwanted mechanical stresses and motion. To limit these effects, devices should be sufficiently removed from the substrate that large z-axis motion cannot be detected. While this definition is fairly loose, it is ultimately up to design and reliability engineers to determine how much parasitic capacitance is tolerable.

The second major source of parasitics comes from within the device. Many new bulk micromachined devices are designed with a silicon base, a reactively grown oxide layer, and a metallization top layer, as shown in Figure 3-14.

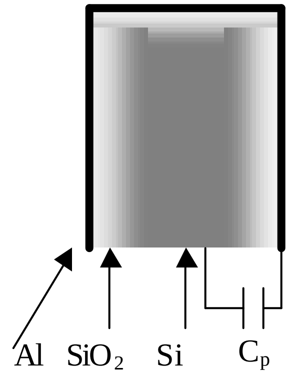

Figure 3-13: Diagram of parasitic capacitance between silicon base and metallization layer on a bulk micromachined beam.

These parasitic capacitances will have varying effects based upon device design. On some devices their effects may be negligible, while on others they can cause severe problems with device output. To limit the influence of parasitic capacitances, many engineers alter their designs to electrically isolate much of their structure. This is done by limiting the amount of metal coated structures and by using thin wires to supply voltages to many structures.

VIII. Dampening Effects

Many MEMS devices are operated in resonant modes, which has some interesting performance implications. All mechanical systems will have specific frequencies at which amplitude, velocity, and acceleration are maximized. They also have an undamped natural frequency, ω_n, which is the oscillating frequency of an unforced system. While in many analyses, this frequency is called the resonant frequency, resonance actually differs from the undamped natural frequency by the relationships below:

$$\text{Displacement resonant frequency,} \; \omega_d = \omega_n \sqrt{\left(1 - 2\zeta^2\right)} \tag{3-21a}$$

$$\text{Velocity resonant frequency,} \; \omega_v = \omega_n \tag{3-21b}$$

$$\text{Acceleration resonant frequency,} \; \omega_a = \frac{\omega_n}{\sqrt{\left(1 - 2\zeta^2\right)}} \tag{3-21c}$$

$$\text{Damped natural frequency,} \; \omega_d = \omega_n \sqrt{\left(1 - \zeta^2\right)} \tag{3-21d}$$

where ζ is the fraction of critical dampening, which is defined as the system dampening, or dampening coefficient, divided by the critical dampening coefficient, c_c. The critical dampening coefficient describes the minimum amount of dampening required for a forced system to return to equilibrium without oscillation, and, for a given system mass, m, and stiffness, k, the mathematical expression for the critical dampening coefficient is given as:[54]

$$c_c = 2\sqrt{km} \tag{3-22}$$

Thus, the resonant frequency is influenced by the system dampening and, if system degradation leads to increased dampening, there will be decreases in resonant frequency.

The typical reason to operate a device at resonance stems from the fact that there is an amplification of system output from the natural response of a structure. The magnitude of this amplification is quantified by the quality factor, Q, which is defined as

$$Q = 2p\frac{W}{\Delta W} \tag{3-23}$$

where W is the energy stored in a system and ΔW is the energy dissipated per cycle. The quality factor also describes the sharpness of the resonance peak. One common method to measure Q is to determine the frequency range for a system at which $v_{out} = v_{max}/\sqrt{2}$, as shown in Figure 3-14. For $\zeta << .1$, the quality factor can be approximated as:

$$\frac{\Delta w}{w_n} = \frac{1}{Q} = 2z \tag{3-24}$$

.

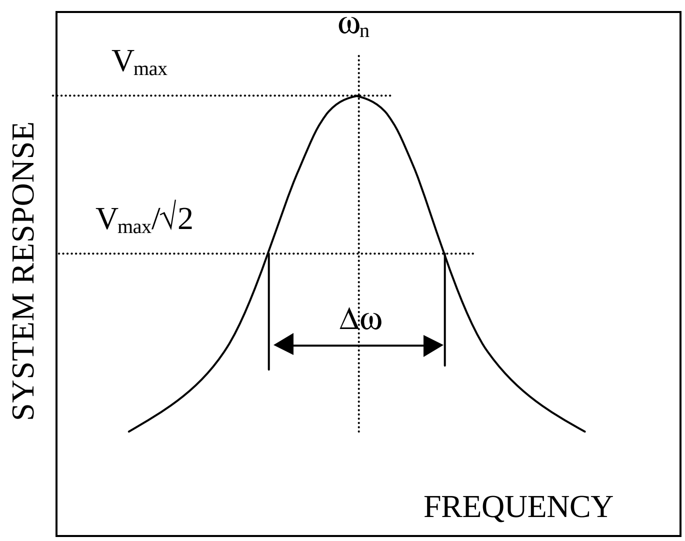

Figure 3-14: Definition of resonance and Q.

As a result, operating a device at resonance, allows greater displacements, which leads to an increase in sensitivity in many systems. However, large structural dampening can cause changes in resonance that will alter output, which can be a long term reliability concern.

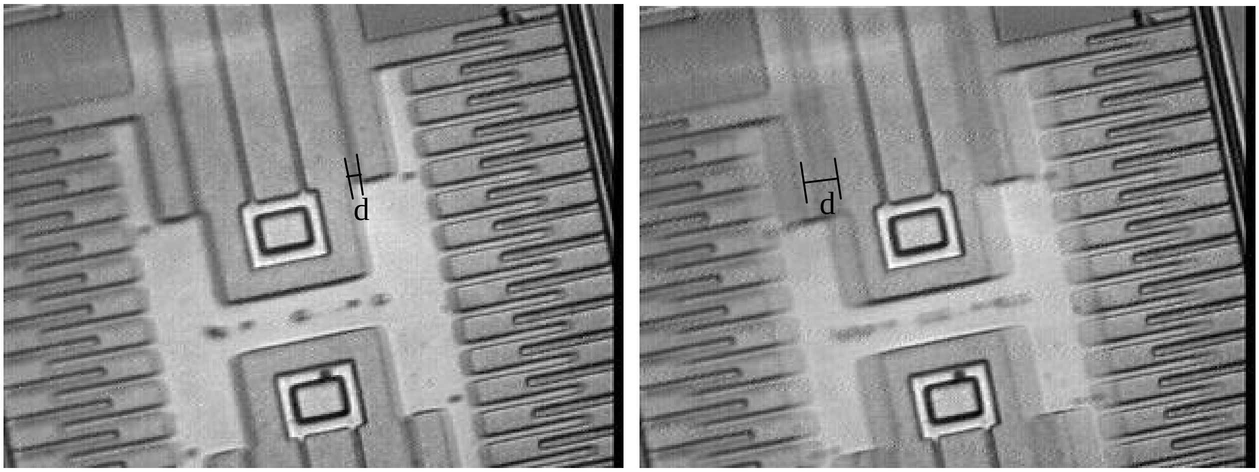

Figure 3-15 (a,b): The difference in amplitude between a linear resonator operating in resonance, right, and operating at a non-resonant frequency, left. As can be seen, the device on the right has a much larger displacement, d, than the device on the left.

MEMS damping is usually caused by the presence of gaseous molecules. There are multiple kinds of damping caused by the atmosphere and the type of damping depends largely upon device geometry. For closely packed parallel surfaces, squeeze film dampening will be predominant. For a rectangular plate of width, 2W, and length, 2L, the squeeze film fraction of critical dampening is, for small displacements[1]:[135,136]

$$z = \frac{8 m f (W/L) W^3 L}{h_0^3 \sqrt{mk}} \qquad (3\text{-}25)$$

where:

f(W/L) = a function of aspect ratio

m = mass of the moving surface

h_0 = distance between the two surfaces at rest

μ = absolute viscosity of air (1.8×10^{-5} Ns/m^2 at 1 ATM)

---

[1] The derivation of this equation assumes that $\omega h^2 \times$(fluid density)/μ <<1.0 .

For a device moving in plane, with an area A, and a separation from the substrate, h, there will be a structural dampening factor due to Couette flow of:[24]

$$z = \frac{mA}{2h\sqrt{mk}} \tag{3-26}$$

Since the dampening is proportional to the viscosity of air, which is a function of pressure, some MEMS utilize vacuum packaging to increase performance. The degree to which a package holds a seal will determine the operating characteristics of these MEMS. Thus changes that lead to increased dampening of a system will alter output by shifting resonant frequency and lowering the quality factor.

IX. Additional Reading

H. G. Van Bueren, Imperfections In Crystals, North-Holland Publishing Company, Amsterdam, 1960.

E. Rabinowicz, Friction and Wear of Materials: 2nd Edition, New York, John Wiley & Sons, 1995.

A. T. DiBenedetto, The Structure and Properties of Materials, McGraw-Hill, United States of America, 1967.